

# 내용기반 영상 검색을 위한 실시간 얼굴 영역 추출

이 미숙, 이 성환

고려대학교 컴퓨터학과 영상처리 연구실  
서울시 성북구 안암동 5가 1번지 (136-701)  
전화 : (02)920-2168, Fax : (02)926-2168  
e-mail : {mslee, swlee}@human.korea.ac.kr

## Real-time Face Extraction for Content-based Image Retrieval

Mee-Sook Lee, Seong-Whan Lee

Image Processing Lab. Dep. of Computer Science and Engineering  
Korea University  
Anam-dong, Seongbuk-ku, Seoul 136-701, Korea

### 요 약

객체 인식은 대용량의 영상 데이터를 분석, 탐색하고 재구성하기 위한 내용기반 영상 검색의 매우 중요한 분야이며, 특히 인간의 얼굴은 검색 영상 내에서 대부분 주요한 장면에 위치하고 있기 때문에 그 비중이 매우 크다. 본 논문에서는 내용기반 영상 검색을 위한 실시간 얼굴 영역 추출 방법을 제안한다. 제안된 방법에서는 다층 피라미드 구조와 간단한 형태의 머리 형판을 사용하여 얼굴의 후보 영역을 추출한 후, 보다 정확한 얼굴 영역을 추출하기 위하여 후보 영역 내에서 눈의 위치를 탐색하고, 두 눈의 위치를 기준으로 최종적인 얼굴 영역을 추출하였다. 얼굴 후보 영역 추출 단계에서는 얼굴의 형태 정보를 포함하고 있는 모자이크 형판을 사용하여 머리와 턱을 포함한 얼굴 영역을 추출하였으며, 눈 위치 추출 단계에서는 눈의 위치 정보를 사용하여 눈의 탐색 영역을 결정하고, 탐색 영역 내에서 이진 영상 형판을 사용하여 눈의 위치를 추출한 후, 눈 영역의 무게 중심을 눈의 중심 위치로 설정하였다. 마지막 얼굴 영역 추출 단계에서는 두 눈의 위치를 기준으로 사각형의 영역을 얼굴 영역으로 추출하였다. 제안된 방법의 성능을 검증하기 위하여 1700장의 다양한 영상에 대하여 실험하였으며, 실험 결과 한 장의 영상에서 얼굴 영역을 추출하는데 있어서, Pentium 166Mz의 PC상에서 평균 3.2초의 처리 속도와 91.7%의 추출률을 보임으로써, 실시간 얼굴 영역 추출에 매우 효과적임을 알 수 있었다.

### 1. 서론

최근 들어 하드웨어 기술의 비약적인 발전으로 텍스트 뿐만이 아니라 영상, 비디오, 오디오 등 다양한 형태의 대용량 멀티미디어 정보를 저장, 검색, 처리할 수 있게 되었으며, 이에 따라 주문형 비디오, 디지털 도서관, 기업의 의사 결정 시스템, 의료 진단 시스템, 각종 교육용 시스템 등 멀티미디어 정보를 사용한 서비스에 대한 요구가 날로 증가하고 있는 추세이다. 그러나 이

러한 멀티미디어 정보 시스템에서 지금까지 주로 사용되어 왔던 검색 방법은 주제어를 중심으로 하는 텍스트 기반의 검색이었으며, 사용자가 찾고자 하는 대상의 주제어를 사전에 알고 있어야만 검색이 가능했었다[1]. 텍스트 기반의 검색 방법은 각 영상에 대하여 입력 시기에 사용된 텍스트 정보에 전적으로 의존하기 때문에 만약 검색 명령어가 초기에 고려되지 않은 속성을 참조하게 되면 대부분의 경우 검색에 실패하게 된다. 특히 특정 질감이나 모양과 같이 복잡한 형태의 가시적

속성을 지닌 데이터의 경우에는 텍스트만으로 속성을 표현하기가 불가능하기 때문에 시각적인 예를 사용한 검색 방법이 요구된다. 내용기반의 검색 시스템은 대용량의 비디오 데이터를 함축적으로 표현하기 위한 키 프레임을 추출하는 부분과, 추출된 키 프레임을 검색에 알맞은 형태로 변환하기 위하여 특징을 추출하고 주석을 첨가하는 영상 색인 부분의 크게 두 부분으로 나눌 수 있다. 이 중에서 키 프레임 추출에 관한 연구는 칼라 히스토그램을 사용한 방법[2], Discrete Cosine 변환을 사용한 방법[3], 다해상도 히스토그램을 사용한 방법 등 활발하게 연구되고 있으나, 객체 인식을 사용한 영상 색인에 관한 연구는 아직도 미미한 상태로 대개는 사람의 손의 통한 수작업으로 진행되고 있다.

따라서 본 논문에서는 내용기반 영상 검색을 위하여 영상 내에 위치한 인간의 얼굴을 자동으로 추출하는 방법에 대하여 기술한다. 본 논문에서 특히 인간의 얼굴을 대상으로 하는 이유는 인간의 얼굴이 영상 내에서 대부분 주요한 장면에 위치하고 있기 때문이며, 뉴스, 스포츠, 다큐멘터리 프로그램의 자동 분류 및 관리 시스템이나 주문형 비디오의 검색 시스템 등 다양한 응용 분야에서 중요한 색인 요소로 사용될 수 있기 때문이다.

제안된 방법에서는 복잡한 배경의 2차원 명도 영상으로부터 얼굴 영역을 추출하기 위하여 얼굴 후보 영역 추출 단계, 눈 위치 추출 단계, 그리고 얼굴 영역 추출 단계의 세가지 단계를 적용하였다. 얼굴 후보 영역 추출 단계에서는 모자이크 형판을 사용한 다층 피라미드 구조의 탐색 방법을 적용하였으며, 다양한 조명 상태에 적응하기 위하여 부분적 히스토그램 균일화를 수행하였다. 눈 위치 추출 단계에서는 추출된 머리 영역 영상에 대해 크기 정규화를 수행하고, 눈에 대한 위치 정보를 이용하여 탐색 영역을 설정하고 이를 이진 영상으로 변환한 후, 눈의 이진 영상 형판을 사용하여 두 눈의 위치를 추출하였다. 그리고 두 눈의 중심 위치를 기준으로하여 눈썹과 입을 포함하는 정사각형의 영역을 얼굴 영역으로 설정하였다.

본 논문의 구성은 다음과 같다. 2장에서는 인간의 얼굴을 사용한 영상 색인과 얼굴 영역 추출에 관한 기존의 연구 동향에 대해서 살펴보고, 3장에서는 형판 정합에 의한 얼굴 후보 영역 추출 과정을 살펴본다. 4장에서는 눈 위치의 추출 과정과 이를 기준으로한 얼굴 영역의 설정 과정에 대해서 설명하고, 5장에서는 실험 및 결과에 대해서 기술한다. 마지막으로 6장에서는 결론 및 앞으로의 연구 방향에 대해 논의한다.

## 2. 관련 연구

내용기반 검색을 위한 영상 색인 기술에서 다루어지는 주요 문제점은 추출된 색인의 정확성과 색인 속도에 관한 것이다. 즉 추출된 영상의 특징이 해당 영상을 정확하게 표현하고, 동시에 다양한 검색에서 사용될 수

있는 것이어야 하며, 실시간 검색을 위해서는 추출 속도가 실시간으로 이루어져야 한다는 것이다. 그러나 영상 색인에 관한 대부분의 연구에서는 실시간 추출을 고려하여 영상의 색깔이나, 질감 등을 사용하여 영상을 표현하였을 뿐이고, 영상 내에 존재하는 객체를 인식하여 표현하려는 연구는 상대적으로 미약한 실정이다.

Pentland[4] 등은 2차원 명도 영상에서 주성분 분석을 사용하여 인간의 얼굴을 표현하는 방법을 제안하고 이를 응용한 "Face Photobook"이라는 도구를 소개하였다. 그러나 이 방법에서 사용한 영상은 얼굴만을 포함한 단순한 형태의 영상이기 때문에 인간의 얼굴을 추출하는 문제는 다루어지지 않았다. Zhang[5]은 지식을 사용한 영상 색인 도구를 소개하면서, 그 응용의 하나로 아나운서 인식을 사용한 뉴스 색인 문제를 소개하였다. 그러나 그가 소개한 방법은 아나운서의 위치가 일반적으로 고정되어 있다는 지식을 기반으로 하고 있기 때문에, 뉴스 이외의 다른 프로그램에는 응용할 수 없다는 단점을 가지고 있다.

2차원 영상에서 얼굴 영역을 추출하는 방법에 대한 연구는 영상 색인에 대한 연구에 비해서 훨씬 일찍부터 연구되어 왔으며, 그 방법에 있어서도 얼굴의 지형적 특징을 사용하는 방법, 통계적 분석법, 특징을 기반으로 한 방법 등 매우 다양하게 연구되어 왔다[6]. 그러나 일반적인 비디오 영상과 같이 복잡한 배경의 영상에서 얼굴을 추출하는 방법에 관해서는 1990년대에 들어서야 본격적으로 연구되기 시작했으며, 실시간 얼굴 영역 추출에 관한 연구는 1990년대 중반에 들어서야 비로소 진행되기 시작했다.

Yang과 Huang[7]들은 배경이 있는 영상 내에서 모자이크 영상을 사용하여 얼굴 영역을 추출하는 방법을 제안하였다. 이 방법에서는 얼굴의 크기 변화에 적응하기 위해서 영상의 해상도를 점차 감소시키면서 특징 정합을 시도하였는데, 60개의 실험 영상에 대하여 적용한 결과 83%의 추출률을 보였다. Kosugi[8]는 Yang과 Huang의 방법을 개선하여 실시간 얼굴 인식에 적합한 얼굴 영역 추출 방법을 제안하였다. 그는 4x5 모자이크 크기의 머리 형판을 사용하여 머리 영역을 추출하고, 추출된 머리 영역 내에서 8x3 모자이크 크기의 눈 형판을 사용하여 얼굴의 중심 위치를 추출한 후, 이를 기준으로 얼굴 영역을 추출하였다. 이 방법은 100개의 실험 영상에 대하여 97%의 추출률을 보였다.

Turk[9] 등은 주성분 분석을 사용하여 얼굴 영역을 추출하였다. 이 방법에서는 주어진 학습 데이터를 사용하여 얼굴 공간(face space)을 구성하고, 입력 영상을 얼굴 공간에 직교 투영한 후, 투영된 영상과 얼굴 공간과의 거리값이 임계치 이하인 영역을 얼굴 영역으로 설정하였다. 기존의 방법들이 소수의 실험 데이터를 사용한 데 비해, 이 방법에서는 약 2500개의 대용량 얼굴 영상을 실험에 사용하여 크기 변화와 방향 변화가 없는

얼굴 영상에 대하여 96%의 추출률을 보였다. 그러나 이 방법은 크기 변화가 가해진 경우 추출률이 60%로 저하되는 등, 방향이나 크기 변화에 매우 민감하다는 단점을 가지고 있다.

국내에서는 김 정환 [10] 등이 다층 퍼셉트론을 사용하여 얼굴 영역을 분할하는 방법에 대하여 제안하였다. 그들은 모자이크 영상을 사용하여 신경망을 학습시키고 영상의 해상도를 변화시켜가면서 얼굴 영역을 추출하였으며, 칼라 정보를 사용하여 탐색 영역을 축소시켜 처리 속도를 증가시켰다. 이 방법은 125개의 실험 영상에 대하여 87%의 추출률을 보였다. 윤 호섭[11] 등은 고립 영역 분석에 의한 얼굴 영역 추출 방법을 제안하였다. 그들은 고립 영역에 대한 사전 지식 및 얼굴 모델 정보를 사용하여 얼굴 영역을 추출하였는데, 예지 영상을 사용하여 눈의 위치를 추출하고 눈을 기준으로 코와 입의 영역을 추출하였다. 이 방법은 80개의 실험 영상에 대하여 91%의 추출률을 보였다.

### 3. 얼굴 후보 영역의 추출

제안된 방법에서는 얼굴 후보 영역을 추출하기 위하여 모자이크 형판을 이용한 형판 정합 방법을 사용하였다. 모자이크 형판을 사용하여 얼굴 후보 영역을 추출하기 위해서는 머리 영역에 대한 정확한 형판 구성이 선행되어야 한다. 본 논문에서 머리 형판의 구성에 사용한 영상들은 비교적 기울어짐이 적고, 편안한 표정의 정면 얼굴 영상으로써, 머리와 턱을 포함한 사각형의 영역을 수동으로 추출하여 히스토그램 균일화를 수행한 후, 모자이크 영상으로 구성하였다. 그리고 성별, 연령 등에 따라 다양한 머리 모양을 자동으로 분류하기 위하여 K-평균 군집화 방법을 사용하였다.

얼굴 후보 영역 추출 단계에서는 다층 피라미드 구조의 탐색 방법을 적용하여 얼굴의 크기 변화를 흡수하였으며, 탐색 시간을 단축하기 위하여 그림 1과 같이 서로 다른 해상도를 갖는 두 종류의 모자이크 형판을 사용한 두 단계의 탐색 방법을 적용하였다.

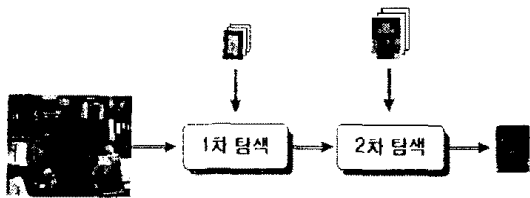


그림 1: 얼굴 후보 영역의 추출 과정

1차 탐색에서는 개략적인 형태의 모자이크 형판을 사용하여 머리의 대략적인 위치를 결정하고, 2차 탐색에서는 상세 정보를 지닌 모자이크 형판을 사용하여 정확한 얼굴 후보 영역을 추출한다. 또한, 조명 및 그림자, 배경 등에 의한 명도 변화에 적응하기 위하여 부분적 히스토그램 균일화를 수행하였으며, 모자이크 형판과의 유사도 비교에는 최근접 이웃 정합 방법을 사용하였다.

### 3.1 머리 형판의 구성

머리 형판 구성에 사용된 데이터는 75명의 남,녀에 대한 기울어짐이 적고, 편안한 표정을 한 얼굴 명도 영상으로써, 머리와 턱을 포함한 사각형의 영역을 수동으로 추출하여 히스토그램 균일화를 수행한 후, 모자이크 영상으로 변환하고, K-평균 군집화를 적용하여 5종류의 머리 형판으로 구성하였다.



(a) 개략 머리 형판



(b) 상세 머리 형판

그림 2: K-평균 군집화에 의해 생성된 머리 형판

그림 2는 K-평균 군집화에 의해서 생성된 머리 형판들로서 다양한 머리 형태에 대한 형판이 구성되었음을 보여주고 있다. 그림 2 (a)의 개략 머리 형판은 머리의 개략적인 형태 정보만을 포함하고 있는 반면, (b)의 상세 머리 형판은 눈과 코, 입의 위치 등, 얼굴의 구성 요소에 대한 정보까지도 포함하고 있음을 알 수 있다.

### 3.2 부분적 히스토그램 균일화

영상의 명도 변화에 적응하기 위한 일반적인 방법으로는 히스토그램 균일화나 히스토그램 신장 등을 사용하게 된다. 그러나 이와 같은 영상 처리 방법을 수행한다고 하더라도 피부색이 검거나 그림자가 드리워진 얼굴 등의 경우에는 주변 배경과의 구분이 모호해진다. 더구나 주변 배경이 피부색보다 밝은 경우 영상 전체에 대해서 히스토그램 균일화를 수행하면 얼굴 영역의 명도 분포가 적용 전보다 어두워지게 되므로 머리 영역 추출이 어려워지게 된다. 따라서 본 논문에서는 탐색 영역 내에서만 히스토그램 균일화를 적용하므로써, 조명 및 피부색, 그림자에 의한 명도 변화에 적응할 수 있도록 하였다.

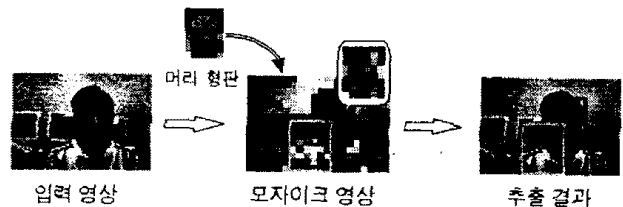


그림 3: 부분적 히스토그램 균일화

그림 3은 부분적 히스토그램 균일화의 예로써, 점선 영역은 영상 전체에 히스토그램 균일화를 적용했을 경우에 대한 결과이고, 실선 영역은 부분적 히스토그램 균일화를 적용했을 경우에 대한 결과이다. 영상의 배경이 얼굴 영역보다 밝고, 얼굴에 그림자가 드리워 있기

때문에 얼굴 영역에 대한 모자이크 영상이 머리 형판에 비하여 어두운데도 불구하고 부분적 히스토그램 균일화를 적용했을 경우에는 머리 영역이 정확하게 추출되었다.

### 3.3 다층 피라미드 구조를 사용한 얼굴 후보 영역의 추출

일반적으로 얼굴 영역 추출에 사용되는 영상에서는 얼굴의 크기가 매우 다양하기 때문에 그림 4과 같이 다층 피라미드 구조를 사용하여 영상의 해상도를 변화시켜 가면서 형판 정합을 시도하게 된다.[8]

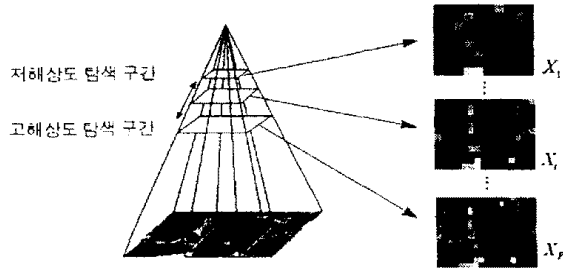


그림 4: 다층 피라미드 구조

그러나 다층 피라미드 구조를 사용한 탐색 방법은 저해상도의 영상에서부터 고해상도의 영상에 이르기까지 다단계의 탐색 과정을 거쳐야 하기 때문에 탐색 시간이 오래 걸린다는 단점을 가지고 있다. 본 논문에서는 추출률에는 커다란 저하없이 탐색 시간을 단축하기 위하여 그림 1과 같이 서로 다른 해상도의 모자이크 형판을 사용한 두 단계의 탐색 과정을 적용하였다. 개략 머리 형판을 사용한 1차 탐색에서는 얼굴의 대략적인 위치를 결정하고, 2차 탐색에서는 상세 머리 형판을 사용하여 탐색 영역 내에서 얼굴의 정확한 위치를 추출하였다. 이와 같은 두 단계의 탐색 과정을 적용함으로써 추출률에는 커다란 저하없이 처리 시간을 3초 정도로 단축시킬 수 있었다.

### 3.4 유사도의 계산

얼굴 후보 영역 추출 단계에서는 탐색 영역 내 영상과 모자이크 형판 사이의 유사도를 계산하기 위해서 최근접 이웃 정합 방법을 사용하였다. 최근접 이웃 정합은 비교되는 두 벡터 사이의 각을 기준으로 계산하기 때문에 명도 변화에는 영향을 적게 받으면서 형태상의 유사도 계산이 가능하다는 장점을 갖는다. 다음은 최근접 이웃 정합 방법에 의한 유사도의 계산 과정이다.

- 각 단계별 모자이크 특징을 벡터 형태로 변환한다.

$$X_i = \{U_{i1}, U_{i2}, \dots, U_{iN_i}\} \quad (1)$$

( $i = 1, 2, \dots, P$  P는 다층 피라미드의 단계)

- 입력 벡터와 머리 형판과의 유사도를 계산한다.

$$S_{ik} = \max\left(\frac{U_{ik} \cdot M_F^j}{\|U_{ik}\| \cdot \|M_F^j\|}\right) \quad (2)$$

( $k = 1, 2, \dots, N_i$   $N_i$ 는  $X_i$ 가 포함한  $U_{ik}$ 의 갯수)

- 각 단계별 최대 유사도를 계산한다.

$$S_i = \max(S_{ik}) \quad (3)$$

- 피라미드 전 단계에 대한 최대 유사도를 계산한다.

$$S = \max(S_i) \quad (4)$$

- 추출된 얼굴 후보 영역에 대한 검증을 수행한다.

$$\text{if}(S \leq T_{Head}) \text{ then reject} \quad (5)$$

식 1에서  $X_i$ 는  $i$ 번째 단계의 모자이크 영상이고,  $U_{ik}$ 는  $X_i$ 를 구성하는  $k$ 번째 벡터로 탐색 영역 내의 모자이크 영상을 의미한다. 식 2에서  $M_F^j$ 는  $j$ 번째 머리 형판을 나타내며, 따라서  $S_{ik}$ 는  $X_i$ 의  $k$ 번째 특징 벡터와의 유사도를 나타낸다. 식 5에서  $T_{Head}$ 는 얼굴 후보 영역에 대한 임계치를 나타내고 있다.

## 4. 얼굴 영역의 추출

모자이크 형판 정합을 사용하여 추출된 얼굴 후보 영역은 배경을 포함하고 있거나, 한 쪽으로 치우치는 등의 위치 오류를 포함하고 있기 때문에, 보다 정확한 얼굴 영역의 추출이 요구된다. 얼굴에 위치한 주요한 특징들은 눈을 기준으로 하여 얼굴의 하단부에 위치해 있으므로, 두 눈의 위치를 기준으로하여 사각형의 얼굴 영역을 추출하였다.

### 4.1 눈의 위치 추출

본 논문에서는 눈의 위치를 추출하기 위하여 구현이 간단하고 처리 시간이 짧은 형판 정합 방법을 사용하였다. 일반적으로 형판 정합에 의한 눈의 위치 추출 방법은 배경 등에 의한 영향을 많이 받으나, 얼굴 후보 영역 추출 과정에서 추출된 영상은 얼굴을 포함한 얼굴 근방의 영역으로 한정되어 있기 때문에 눈의 위치 정보를 사용하여 이에 따르는 오류를 방지할 수 있었다. 그림 5는 눈의 위치 추출 과정에 대한 예이다.



그림 5: 눈의 위치 추출 과정

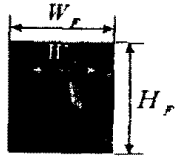
### 4.2 얼굴 영역의 추출

사람은 일반적으로 정자세를 취하고 있는 경우에도 얼굴이 약간씩은 기울어지는 경향이 있기 때문에 이를 수평으로 교정해 주어야만 한다. 따라서 얼굴 영역 추출 단계에서는 식 6과 같은 회전 변환식을 사용하여 얼굴의 기울기를 보정하였다.

$$\begin{vmatrix} x' \\ y' \end{vmatrix} = \begin{vmatrix} \cos \theta & \sin \theta \\ \sin \theta & \cos \theta \end{vmatrix} \begin{vmatrix} x \\ y \end{vmatrix} \quad (6)$$

식 6에서  $(x, y)$ 는 회전 변환한 영상에서 두 눈의 중심 위치를 기준으로 한 상대 좌표이며,  $(x', y')$ 는  $(x, y)$ 에 대응하는 회전 전 영상 내에서의 좌표이다.

얼굴 영역은 그림 6에서와 같이 두 눈의 위치를 중심으로 눈썹과 턱을 포함하는 정사각형의 영역으로 설정하였다.



$$H_F = W_E \times 2, W_F = W_E \times 2$$

그림 6: 얼굴 영역의 범위

## 5. 실험 및 결과 분석

### 5.1 실험 환경

제안된 방법의 성능의 검증에 위한 실험에서는 인간의 얼굴이 포함되어 있는 얼굴 영상과 포함되어 있지 않은 일반 영상, 두 종류에 대하여 총 1700장을 사용하였다. 얼굴 영상은 Pentium 166MHz의 일반 PC 상에서 CCD 카메라로 획득한 320x240, 240x220 해상도의 컬러 영상을 사용하였으며, 일반 영상은 인터넷 상에 공개된 영상을 사용하였다. 얼굴 영상으로 사용된 데이터는 다양한 연령층의 남, 녀 75명에 대한 영상으로써, 동일한 사람에 대하여 배경과 조명 상태, 카메라와의 거리 등을 변화시키면서 획득하였다. 그림 7은 제안된 방법의 성능을 검증하기 위해 사용된 영상들에 대한 예이다.



그림 7: 얼굴 영역 추출에 사용된 영상

### 5.2 실험 결과

표 1은 사용된 머리 형판의 종류에 따른 처리 속도와 추출률을 비교한 것으로서, 4x5 형판이나 6x8 형판이 8x10 형판에 비해서 추출률이 더 높다는 것을 알 수 있다. 이는 8x10 형판에는 추출에 필요한 정보 이외의

불필요한 정보가 추가적으로 더 포함되어 있기 때문에 나타나는 현상이다.

표 1: 머리 형판의 종류에 따른 실험 결과

형판의 종류	4x5	6x8	8x10
시간(초)	3.8	6.2	7.5
추출률	96.0%	96.6%	91.3%

표 2: 추출방법에 따른 실험 결과

추출 방법	4x5(일차)	6x8	6x8(일차)
	6x8(이차)	(임계치)	6x8(이차)
시간(초)	3.2	6.4	5.5
추출률	94.6%	96.0%	93.3%

표 2는 추출 방법에 따른 실험 결과로써, (일차)는 일차 탐색, (이차)는 이차 탐색, (임계치)는 유사도가 임계치 이상인 영역을 얼굴 후보 영역으로 설정한 경우를 나타내고 있다. 표 2에 따르면 6x8 형판을 사용한 한번의 탐색 방법이 추출률은 가장 좋으나 처리 시간이 6초가 소요되는 반면, 제안된 방법이 추출률은 94.6%로 6x8 형판의 경우에 비해서 다소 떨어지나 처리 속도는 3초로써 실시간 얼굴 후보 영역 추출에는 더 효과적임을 알 수 있었다.

눈의 위치 추출 실험은 얼굴 후보 영역 추출에 성공한 영상에 대해서만 적용하여, 97.1%의 추출률을 보였으며, 얼굴 영역의 추출은 단순히 눈을 기준으로 사각형의 영역을 설정하는 것이므로, 전체 영상에 대한 얼굴 영역의 추출률은 얼굴 영역의 추출률과 눈의 추출률과의 곱으로 계산될 수 있으며, 91.7%가 된다.

그림 8은 얼굴 영역 추출에 성공한 예로써 얼굴의 피부색이 비교적 검은 편이거나 그림자가 드리워진 경우, 얼굴이 약간 기울어진 경우에도 정확한 추출이 가능하였다.

그림 9은 얼굴 영역 추출에 실패한 예로써 첫번째 영상의 경우에는 검은색 이외의 머리를 가지고 있기 때문에 형판 정합에서 실패한 경우이고, 두 번째와 세 번째 영상은 얼굴의 방향이 정면에서 많이 벗어났기 때문에 실패한 경우를 보여주고 있다.



그림 8: 얼굴 영역 추출에 성공한 예



그림 9: 얼굴 영역 추출에 실패한 예

표 3은 제안된 방법을 기존의 얼굴 영역 추출 방법들과 비교한 결과를 보여주고 있다. 비교 대상 방법으로는 Turk와 Pentland가 제안한 PCA를 사용한 방법, Yang과 Huang이 제안한 얼굴 영역 모자이크를 사용한 방법, 그리고 Kosugi가 제안한 다해상도 영상을 사용한 방법을 선택하였다. 결과적으로 Kosugi의 방법과 제안된 방법이 처리 시간과 추출률면에서 다른 두 가지 방법보다 우수하다는 것을 알 수 있다.

표 3: 기존 방법과의 비교

추출 방법	PCA	얼굴영역 모자이크	다해상도 영상	제안된 방법
시간(초)	7.48	12.27	4.80	3.20
추출률(%)	76.6	78.0	93.8	94.6

## 6. 결론 및 향후 연구 방향

본 논문에서는 내용기반 영상 검색을 위한 영상 색인 기술의 하나로써 영상 내에서 실시간으로 얼굴 영역을 추출하는 방법에 대하여 제안하였다.

복잡한 배경을 가진 영상으로부터 다양한 크기의 얼굴을 추출하기 위하여 모자이크 형판을 사용한 다층 피라미드 구조의 탐색 방법을 적용하였으며, 조명 및 피부색, 그림자 등에 의한 명도 변화에 적응하기 위하여 부분적 히스토그램 균일화 방법을 사용하였다. 또한 서로 다른 해상도의 머리 형판을 사용한 두 단계의 탐색 방법으로 추출률에는 큰 저하없이 처리 시간을 3초로 단축시킬 수 있었다. 눈 위치의 추출 단계에서는 눈에 대한 위치 정보를 사용하여 탐색 영역을 설정하고, 눈의 이진 영상 형판을 사용하여 눈의 중심 위치를 추출하였고, 두 눈의 중심 위치를 기준으로 얼굴 영역을 추출하였다.

제안된 방법의 성능을 검증하기 위해서 1700장의 다양한 실험 영상에 대하여 적용한 결과, 영상 한 장당 얼굴 영역 추출에 걸리는 처리 시간은 평균 3.2초, 추출률은 91.7%을 보여 실시간 얼굴 영역 추출에 효과적임을 보였다. 그러나 제안된 방법은 대머리인 경우에는 추출에 실패하는 단점을 가지고 있다. 따라서 다양한 머리 형태에 대한 머리 형판의 구성이 필요하며, 보다 정확한 얼굴 영역 추출을 위해서 동영상을 사용한 방법에 대한 연구가 수행되어야 한다. 또한 영상 내에 다수의 얼굴이 존재할 수 있으므로 이에 대한 해결 방안도 연구되어야 할 것이다. 그리고 보다 효과적인 영상 색인을 위해서는 얼굴 영역 추출 뿐만이 아니라 얼굴 인식 기술에 대한 연구도 수행되어야 할 것이다.

## 참고 문헌

[1] 정 원일, 최 현섭, 최 기호. "내용기반 화상 검색 시

스템의 설계 및 구현," 전자공학회논문지 제 33권 B편 제 7호, pp. 1080-1089, 1996년 7월.

- [2] Y. Gong, C. H. Chuan and G. Xiaoyi, "Image Indexing and Retrieval Based on Color Histograms," *Multimedia Tools and Applications*, Vol. 2, pp. 133-156, 1996.
- [3] V. V. Vinod and H. Murase, "Object Location Using Complementary Color Features: Histogram and DCT," *Proc of ICPR'96*, Vol.I, pp. 554-559, 1996.
- [4] A. Pentland, R. W. Picard, and S. Sclaroff, "Photobook: Tools for Content-Based Manipulation of Image Databases," *Proc. of SPIE'94*, Vol. 2185, pp. 34-47, 1994.
- [5] H. Zhang and S. W. Smoliar, "Developing Power Tools for Video Indexing and Retrieval," *Proc. of SPIE'94*, Vol. 2185, pp. 140-149, 1994.
- [6] 이 성환, 이 미숙, "얼굴 영상 인식 기술의 현황," 대한전자공학회지, 제 23권, 제 6호, pp. 80-94, 1996년 7월.
- [7] G. Yang and T. S. Huang, "Human Face Detection in a Complex Background," *Pattern Recognition*, Vol. 27, No. 1, pp. 53-63, 1994.
- [8] M. Kosugi, "Human-Face Search and Location in a Scene by Multipyramid Architecture for Personal Identification," *Systems and Computers in Japan*, Vol. 26, No. 6, pp. 27-38, 1995.
- [9] M. A. Turk and A. P. Pentland, "Face Recognition Using Eigenfaces," *Proc. of CVPR'91*, pp. 586-591, 1991.
- [10] 김 정환, 박 강령, 고 희동, 김 재희, "다층 퍼셉트론을 이용한 얼굴 영역 분할 및 특징부 추출," '96 FAN 인공지능, 신경망 및 퍼지 시스템 학술발표회 논문집, pp. 159-165, 1996.
- [11] 윤 호섭, 소 정, 왕 민, 민 병우, "고립 영역 분석에 의한 얼굴 요소 추출," 한국정보과학회 논문지(B), 제 23권, 제 7호, pp. 752-765, 1996.